AD-A060 236 NAVY UNDERWATER SOUND LAB NEW LONDON CONN F/G 17/1
ON THE EFFECT OF A BIAS ERROR IN THE ADAPTIVE LMS ALGORITHM.(U)
APR 70 N L OWSLEY
UNCLASSIFIED NUSL-TM-2242-87-70 NL

1 OF 1
AD A060236

END
DATE
FILMED
12-78
DDC

Code No. ____

Copy No. 3

NAVY UNDERWATER SOUND LABORATORY
FORT TRUMBULL, NEW LONDON, CONNECTICUT

6

ON THE EFFECT OF A BIAS ERROR IN THE ADAPTIVE LMS ALGORITHM,

by

10 Norman L./Owsley

NUSL Technical Memorandum No. 2242-87-70

11 7 April 1970

12 8 p.

Introduction

9 Technical memo,

A basic assumption in the derivation of the adaptive beamformer LMS algorithm is that the input sample vector is of zero mean.[1] In the actual hardware implementation of this algorithm, bias errors can be encountered due to system electronics. For example, if the beamformer is operated for a considerable length of time, then a small bias in a nominally even A/D quantizer can seemingly cause erroneous algorithm convergence with respect to the expected zero mean system input. This memorandum considers some of the elemental aspects of such biasing errors.

The Biased LMS Algorithm

Consider the sampled vector

$$\underline{X}_B = \underline{X} + \underline{B} \quad (1)$$

14 NUSL-TM-2242-87-70

where $\underline{X}$ is the true, zero mean system input vector and $\underline{B}$ is a constant vector which is introduced to account for biasing of the actual data. Therefore, if the superscript "T" represents the vector transpose, then

$$E\left\{\underline{X}_B\right\} = \underline{B} \quad (2)$$

254 200

and the correlation matrix R_{BB} is given by

$$R_{BB} = E\left\{\underline{X}_B \underline{X}_B^T\right\} \tag{3}$$

$$= R_{XX} + \underline{B}\,\underline{B}^T \tag{4}$$

where $E\{\cdot\}$ is the expectation operator and $\underline{X}$ and $\underline{B}$ are assumed to be uncorrelated.

The beamformer output can be specified as

$$y(t) = \underline{W}^T \underline{X}_B \tag{5}$$

$$= \underline{X}_B^T \underline{W}\,. \tag{6}$$

According to the Least Mean Square (LMS) optimality criterion, the expected squared error

$$E\left\{\varepsilon^2\right\} = E\left\{[y(t) - d(t)]^2\right\} \tag{7}$$

$$= E\left\{[\underline{W}^T \underline{X}_B - d(t)]^2\right\} \tag{8}$$

is to be minimized by selecting $\underline{W}$ to give a best mean square estimate of the desired signal, $d(t)$. Expanding (8) gives

$$E\left\{\varepsilon^2\right\} = \underline{W}^T R_{BB} \underline{W} - 2\underline{W}^T E\left\{\underline{X}_B d(t)\right\} + E\left\{d^2(t)\right\} \tag{9}$$

$$= \underline{W}^T R_{XX} \underline{W} + \underline{W}^T \underline{B}\,\underline{B}^T \underline{W} - 2\underline{W}^T \underline{P} + E\left\{d^2(t)\right\} \tag{10}$$

where $\underline{B}$ and d(t) are assumed to be uncorrelated. The mean squared error (10), is now minimized by taking the gradient of (10) with respect to $\underline{W}$, equating to zero and solving for $\underline{W}_B$. Accordingly,

$$\nabla_{\underline{W}} E\left\{\varepsilon^2\right\} = 2[R_{xx} + \underline{B}\ \underline{B}^T]\underline{W}_B - 2\underline{P}$$

$$= \underline{0}. \tag{11}$$

Solving for $\underline{W}_B$, the biased optimum weight vector, gives

$$\underline{W}_B = [R_{xx} + \underline{B}\ \underline{B}^T]^{-1}\underline{P}. \tag{12}$$

Notice, that for $\underline{B} = \underline{0}$, i.e. zero input bias $\underline{W}_B$ becomes

$$\underline{W}_{LMS} = R_{xx}^{-1}\ \underline{P} \tag{13}$$

the desired LMS weight vector. Thus the bias vector can be visualized as adding a constant bias power term to each element of the actual correlation matrix R_{xx}.

Woodbury's identity can be applied to Equation (12) to give

$$\underline{W}_B = \left[R_{xx}^{-1} - \frac{R_{xx}^{-1}\underline{B}\ \underline{B}^T\ R_{xx}^{-1}}{1 + \underline{B}^T R_{xx}^{-1}\ \underline{B}}\right]\underline{P} \tag{14}$$

$$= \left[I - \frac{R_{xx}^{-1}\ \underline{B}\ \underline{B}^T}{1 + \underline{B}^T R_{xx}^{-1}\ \underline{B}}\right]R_{xx}^{-1}\underline{P}. \tag{15}$$

If the same bias component, b, is present in all terms of $\underline{X}_B$ then $\underline{B}$ can be written as

$$\underline{B} = b\,\underline{1} \tag{16}$$

where $\underline{1}$ is a vector of ones with the same dimension as $\underline{X}$. Now, Equation (15) reduces to

$$\underline{W}_B = \left[I - \frac{R_{xx}^{-1}\,\underline{1}\,\underline{1}^T}{\frac{1}{b^2} + \underline{1}^T R_{xx}^{-1}\,\underline{1}} \right] R_{xx}^{-1}\,\underline{P}\,. \tag{17}$$

In addition, if R_{xx} is the correlation matrix for uncorrelated noise of power N_o, and p_i is the i^{th} element of $\underline{P}$, then Equation (17) becomes

$$\underline{W}_B = \frac{1}{N_o} \left[I - \frac{\underline{1}\,\underline{1}^T}{\frac{N_o}{b^2} + KL} \right] \underline{P} \tag{18}$$

$$= \frac{\underline{P}}{N_o} - \frac{1}{N_o}\,\frac{\sum_{i=1}^{KL} p_i}{\frac{N_o}{b^2} + KL}\,\underline{1} \tag{19}$$

where K is the number of channels in the system and L is the number of time delays per channel. Thus, only if

$$N_o\left(\frac{N_o}{b^2} + KL\right) >> \sum_{i=1}^{KL} p_i \qquad (20)$$

the optimum weight vector for this case, namely

$$\underline{W}_{LMS} = \frac{\underline{P}}{N_o}, \qquad (21)$$

would not be corrupted significantly.

Convergence of the Biased LMS Algorithm

The instantaneous gradient of ε^2 with respect to $\underline{W}$ is obtained from (10) as

$$\varepsilon^2(i) = \underline{W}^T(i)\underline{X}(i)\underline{X}^T(i)\underline{W}(i) + \underline{W}^T(i)\underline{B}\underline{B}^T\underline{W}(i) - 2\,\underline{W}(i)\underline{P} + d^2(t_i) \qquad (22)$$

where the $i^{\underline{th}}$ sampling instant is considered. The gradient of $\varepsilon^2(i)$ with respect to $\underline{W}(i)$ is

$$\nabla_{\underline{W}(i)}\,\varepsilon^2(i) = \underline{J}(\underline{W}(i)) \qquad (23)$$

$$= 2[\underline{X}(i)\underline{X}(i) + \underline{B}\,\underline{B}^T]\underline{W}(i) - 2\,\underline{P} \qquad (24)$$

giving the gradient search algorithm

$$\underline{W}(i+1) = \underline{W}(i) + \mu \underline{J}(\underline{W}(i)) \tag{25}$$

or

$$\underline{W}(i+1) = [I + 2\mu(\underline{X}(i)\underline{X}(i) + \underline{B}\ \underline{B}^T)]\underline{W}(i) - 2\mu\underline{P}\ . \tag{26}$$

The mean value of (26) for uncorrelated $\underline{X}(i)$ and $\underline{W}(i)$ is

$$\underline{M}(i+1) = E\left\{W(i+1)\right\} \tag{27}$$

$$= [I + 2\mu\ (R_{xx} + \underline{B}\ \underline{B}^T)]\underline{M}(i) - 2\ \underline{P}. \tag{28}$$

A standard convergence proof would allow convergence of (28) as follows

$$\lim_{n\to\infty} \underline{M}(n) = [R_{xx} + \underline{B}\ \underline{B}^T]^{-1}\underline{P} \tag{29}$$

only if

$$-\frac{1}{2\lambda_{max}} < \mu < o \tag{30}$$

where λ_{max} is the maximum eigenvalue of $(R_{xx} + \underline{B}\ \underline{B}^T)$.

Bias Removal

The system of Figure 1 is suggested for the removal of bias from the input sample vector. The vector $\underline{B}(i)$ is

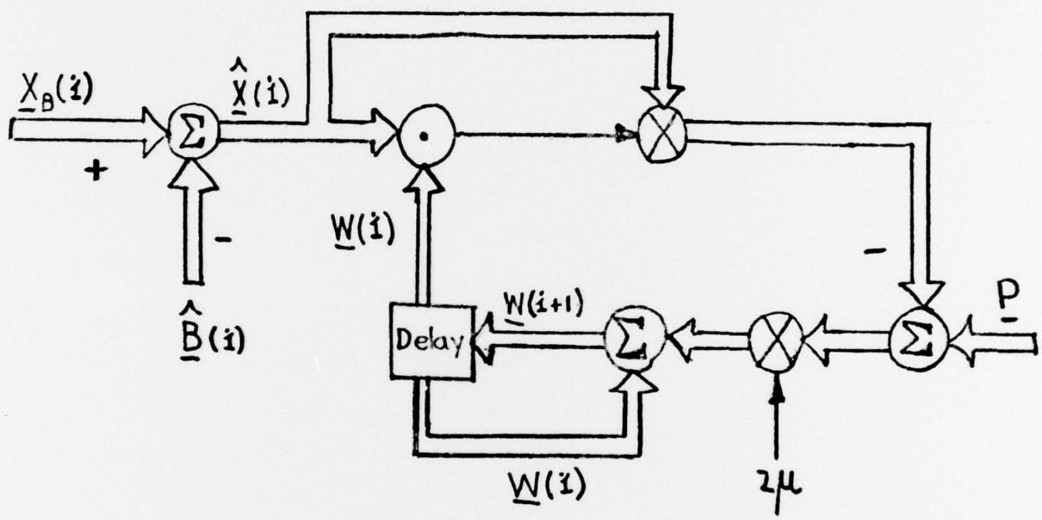

Figure 1. Bias Removal LMS System

an estimate of the actual $\underline{B}$ and could perhaps be obtained by measuring, over an interval, the mean of the biased input sample vector $\underline{X}_B$.

Conclusions

Bias errors due to system electronics for the implementation of the LMS adaptive beamformer have been considered. Basically, the bias contributes to system error by adding a constant term, or bias power, to each term in the actual system input correlation matrix. The result is a biased LMS filter weight vector (see Equation (15)).

To obtain some insight, the simple case of an uncorrelated noise input was considered. It was seen that the optimum weights were all biased by the constant factor. The amount of the bias is proportional to both the system bias power and the sum of the elements in the steering vector, $\underline{P}$.

Also, an LMS system modification is suggested which conceivably could eliminate the bias at the input by subtracting an estimate of the mean input sample vector.

Norman L. Owsley

Norman L. Owsley
Research Electronic Engineer

Reference:

(1) Widrow, B., Montey, P., Griffiths, L., and Goode, B., "Adaptive Antenna Systems", Proceedings of the IEEE, Vol. 55, No. 12, Dec. 1967.